SEQUENCE LISTING

<110> Tang, Y. Tom Zhou, Ping

| Liu, Asun Ren, Zhac Yang Wehr | rich, Ryle Chenghua di, Vinod Feiyan , Qing A. , Yonghong man, Tom hac, Radoje | • Т. | | | | |
|--|--|--------------|------------|------------|------------|-----|
| | Novel Nucle peptides | eic Acids ar | nd | | | |
| <130> | 787CIP2F | | | | | |
| <140> <141> | To Be Assig | gned | | | | |
| | 09/560,875 2000-04-27 | | | | | |
| | 09/496,914 2000-02-03 | | | | | |
| <160> | 10 | | | | | |
| <170> | pt_FL_genes | Version 2 | . 0 | | | |
| | | | | | | |
| <210> <211> <212> <213> | 1313 | ıs | | | | |
| <220> <221> <222> | CDS (547)(123 | 39) | | | | |
| <400> tagatgtgtt | | agccgcaaaa | catttctcag | tcttagtctc | ttgtaaaata | 60 |
| gaggtaataa | gaaacacttt | tcagtatttt | gtgacatgta | gaagtaagtg | atggtggcat | 120 |
| gcatcacact | tggttaatag | taggtcctgt | tgttaagtct | ctaatggcga | taccctatgg | 180 |

240 cttctccaaa tggtgacctt gccaaattgt tttccaaagc gacatgtggc ttttttctcc 300 caatccctca ttttaactct catggtaatt taacttttat atttttatta gatgcattta gtaacttgcc tcatagtcat tttcttggaa attcaatttc ttctccacag ggtctctttt 360 420 qaqattaaaq agaqagaagt ggcaaattta ggatgttaga ataattttca tttaaaagta gatccttgtt tttattaccc tatcattaat gttttctgtt ttcctttatc agcgagttac 480 tgctcatttg attcatattg ccaaactgaa ctctcttgtt ttcttgcaag atgaaaggag 540 588 atg aat gag cca cta gac tat tta gca aat gct tct gat ttc Met Asn Glu Pro Leu Asp Tyr Leu Ala Asn Ala Ser Asp Phe 636 ccc gat tat gca gct gct ttt gga aat tgc act gat gaa aac atc cca Pro Asp Tyr Ala Ala Ala Phe Gly Asn Cys Thr Asp Glu Asn Ile Pro 15 ctc aag atg cac tac ctc cct gtt att tat ggc att atc ttc ctc gtg 684 Leu Lys Met His Tyr Leu Pro Val Ile Tyr Gly Ile Ile Phe Leu Val 35 40 732 gga ttt cca ggc aat gca gta gtg ata tcc act tac att ttc aaa atg Gly Phe Pro Gly Asn Ala Val Val Ile Ser Thr Tyr Ile Phe Lys Met 50 55 aga cct tgg aag agc acc atc att atg ctg aac ctg gcc tgc aca 780 Arg Pro Trp Lys Ser Ser Thr Ile Ile Met Leu Asn Leu Ala Cys Thr 70 65 gat ctg ctg tat ctg acc agc ctc ccc ttc ctg att cac tac tat gcc 828 Asp Leu Leu Tyr Leu Thr Ser Leu Pro Phe Leu Ile His Tyr Tyr Ala 80 85 876 agt ggc gaa aac tgg atc ttt gga gat ttc atg tgt aag ttt atc cgc Ser Gly Glu Asn Trp Ile Phe Gly Asp Phe Met Cys Lys Phe Ile Arg 924 ttc agc ttc cat ttc aac ctg tat agc agc atc ctc ttc ctc acc tgt Phe Ser Phe His Phe Asn Leu Tyr Ser Ser Ile Leu Phe Leu Thr Cys 115 972 ttc agc atc ttc cgc tac tgt gtg atc att cac cca atg agc tgc ttt Phe Ser Ile Phe Arg Tyr Cys Val Ile Ile His Pro Met Ser Cys Phe 135 130 1020 tcc att cac aaa act cga tgt gca gtt gta gcc tgt gct gtg gtg tgg Ser Ile His Lys Thr Arg Cys Ala Val Val Ala Cys Ala Val Val Trp 150 145 atc att tca ctg gta gct gtc att ccg atg acc ttc ttg atc aca tca 1068 Ile Ile Ser Leu Val Ala Val Ile Pro Met Thr Phe Leu Ile Thr Ser 170 160 165

| acc aac agg acc aac aga tca gcc tgt ctc gac ctc acc agt tcg gat Thr Asn Arg Thr Asn Arg Ser Ala Cys Leu Asp Leu Thr Ser Ser Asp 175 180 185 190 | 1116 |
|---|--------------------------|
| gaa ctc aat act att aag tgg tac aac cta att ttg act gca act act Glu Leu Asn Thr Ile Lys Trp Tyr Asn Leu Ile Leu Thr Ala Thr Thr 195 200 205 | 1164 |
| ttg cct ccc ctt ggt gat agt gac act ttg cta tac cac gat tat cca Leu Pro Pro Leu Gly Asp Ser Asp Thr Leu Leu Tyr His Asp Tyr Pro 210 215 220 | 1212 |
| cac tct gac cca tgg act gca aac tga cagct gccttaagca gaaagcacga His Ser Asp Pro Trp Thr Ala Asn * 225 230 | 1264 |
| aggctaacca ttctgctact cccggacgcg tgggtcgaca cgggaatgt | 1313 |
| <210> 2 <211> 2751 | |
| <212> DNA <213> Homo sapiens | |
| <220> <221> CDS <222> (58)(1482) | |
| (222) (30)(1402) | |
| <400> 2 ggcttcgggg tgggcgcca gggtagggca ggtggccgcc gcgtggaggc agggaga | 57 |
| <400> 2 | 57 105 |
| <pre><400> 2 ggcttcgggg tgggcgccca gggtagggca ggtggccgcc gcgtggaggc agggaga atg cga ctc tcc aaa acc ctc gtc gac atg gac atg gcc gac tac agt Met Arg Leu Ser Lys Thr Leu Val Asp Met Asp Met Ala Asp Tyr Ser</pre> | |
| <pre> <400> 2 ggcttcgggg tgggcgccca gggtagggca ggtggccgcc gcgtggaggc agggaga atg cga ctc tcc aaa acc ctc gtc gac atg gac atg gcc gac tac agt Met Arg Leu Ser Lys Thr Leu Val Asp Met Asp Met Ala Asp Tyr Ser 1</pre> | 105 |
| <pre> </pre> <pre> <pre> <a hre<="" td=""><td>105 153</td></pre></pre> | 105 153 |
| <pre> <400> 2 ggcttcgggg tgggcgccca gggtagggca ggtggccgcc gcgtggaggc agggaga atg cga ctc tcc aaa acc ctc gtc gac atg gac atg gcc gac tac agt Met Arg Leu Ser Lys Thr Leu Val Asp Met Asp Met Ala Asp Tyr Ser 1</pre> | 105 153 201 |
| <pre>c400> 2 ggcttcgggg tgggcgccca gggtagggca ggtggccgcc gcgtggaggc agggaga atg cga ctc tcc aaa acc ctc gtc gac atg gac atg gcc gac tac agt Met Arg Leu Ser Lys Thr Leu Val Asp Met Asp Met Ala Asp Tyr Ser 1</pre> | 105 153 201 249 |

| Arg | Phe | Ser | Arg 100 | Gln | Cys | Val | Val | Asp 105 | Lys | Asp | Lys | Arg | Asn 110 | Gln | Cys | |
|-----|-------------------|-----|------------|-----|-----|-----|-----|------------|-----|-----|-----|-----|------------|-----|-----|------|
| | tac Tyr | | | | | | | | | | | | | | | 441 |
| | gtc Val 130 | | | | | | | | | | | | | | | 489 |
| | gac Asp | | | | | | | | | | | | | | | 537 |
| | tcc Ser | | | | | | | | | | | | | | | 585 |
| | gcg Ala | | | | | | | | | | | | | | | 633 |
| | cag Gln | | | | | | | | | | | | | | | 681 |
| | gag Glu 210 | | | | | | | | | | | | | | | 729 |
| | gag Glu | | | | | | | | | | | | | | | 777 |
| _ | gtg Val | _ | | | | | | | | | | | | | | 825 |
| | ctg Leu | | | | | | | | | | | | | | | 873 |
| | ctg Leu | | | | | | | | | | | | | | | 921 |
| | aaa Lys 290 | | | | | | | | | | | | | | | 969 |
| | Gly aaa | | | | | | | | | | | | | | | 1017 |
| | tac Tyr | | | | | | | | | | | | | | | 1065 |

325 330 335

| ctg ctg ctg ctg ccc acc ttg cag agc atc acc tgg cag atg atc Leu Leu Leu Leu Pro Thr Leu Gln Ser Ile Thr Trp Gln Met Ile 340 345 350 | 1113 |
|---|------|
| gag cag atc cag ttc atc aag ctc ttc ggc atg gcc aag att gac aac Glu Gln Ile Gln Phe Ile Lys Leu Phe Gly Met Ala Lys Ile Asp Asn 355 360 365 | 1161 |
| ctg ttg cag gag atg ctg ctg gga ggg tcc ccc agc gat gca ccc cat Leu Leu Gln Glu Met Leu Leu Gly Gly Ser Pro Ser Asp Ala Pro His 370 375 380 | 1209 |
| gcc cac cac ccc ctg cac cct cac ctg atg cag gaa cat atg gga acc Ala His His Pro Leu His Pro His Leu Met Gln Glu His Met Gly Thr 385 390 395 400 | 1257 |
| aac gtc atc gtt gcc aac aca atg ccc act cac ctc agc aac gga cag Asn Val Ile Val Ala Asn Thr Met Pro Thr His Leu Ser Asn Gly Gln 405 410 415 | 1305 |
| atg tgt gag tgg ccc cga ccc agg gga cag gca gcc acc cct gag acc Met Cys Glu Trp Pro Arg Pro Arg Gly Gln Ala Ala Thr Pro Glu Thr 420 425 430 | 1353 |
| cca cag ccc tca ccg cca ggt ggc tca ggg tct gag ccc tat aag ctc Pro Gln Pro Ser Pro Pro Gly Gly Ser Gly Ser Glu Pro Tyr Lys Leu 435 440 445 | 1401 |
| ctg ccg gga gcc gtc gcc aca atc gtc aag ccc ctc tct gcc atc ccc Leu Pro Gly Ala Val Ala Thr Ile Val Lys Pro Leu Ser Ala Ile Pro 450 455 460 | 1449 |
| cag ccg acc atc acc aag cag gaa gtt atc tag caagccgc tggggcttgg Gln Pro Thr Ile Thr Lys Gln Glu Val Ile * 465 470 475 | 1500 |
| gggetecaet ggeteceee ageeeeetaa gagageaeet ggtgateaeg tggteaegge | 1560 |
| aaaggaagac gtgatgccag gaccagtccc agagcaggaa tgggaaggat gaagggcccg | 1620 |
| agaacatggc ctaagggcca catcccactg ccacccttga cgccctgctc tggataacaa | 1680 |
| gactttgact tggggagacc tctactgcct tggacaactt atctcatgtt gaagccactg | 1740 |
| ccttcacctt caccttcatc catgtccaac ccccgacttc atcccaatgg acagccgcct | 1800 |
| ggagatgact tgaggcctta cttaaaccca gctcccttct tccctagcct ggtgcttctc | 1860 |
| ctctcctagc ccctgtcatg gtgtccagac agagccctgt gaggctgggt ccaattgtgg | 1920 |
| cacttggggc accttgctcc tccttctgct gctgccccca cctctgctgc ctccctctgc | 1980 |
| tgtcaccttg ctcagccatc ccgtcttctc caacaccacc tctccagagg ccaaggaggc | 2040 |
| cttggaaacg attcccccag tcattctggg aacatgttgt aagcactgac tgggaccagg | 2100 |

2160 tcatcctcct tcttcaggga cttgggtggg tacttgggtg aggatccctg aaggccttca 2220 acccgagaaa acaaacccag gttggcgact gcaacaggaa cttggagtgg agaggaaaag 2280 catcagaaag aggcagacca tccaccaggc ctttgagaaa gggtagaatt ctggctggta 2340 gagcaggtga gatgggacat tccaaagaac agcctgagcc aaggcctcgt ggtagtaaga 2400 atctagcaag aattgaggaa gaatggtgtg ggagagggat gatgaagaga gagagggcct 2460 actggagage atagggtetg gaacaccagg etgaggteet gateagette aaggagtatg 2520 cagggagctg ggcttccaga aaatgaacac agcagttctg cagaggacgg gaggctggaa 2580 gctgggaggt caggtggggt ggatgatata atgcgggtga gagtaatgag gcttggggct 2640 2700 ggagaggaca agatgggtaa acceteacat cagagtgaca tecaggagga ataagteeca gggcctgtct ctaggtcgtc tgtaacgccg gccatgtcca gcgtcagtat g 2751

<210> 3

<211> 6109

<212> DNA

<213> Homo sapiens

<220>

<221> CDS

<222> (227)..(2836)

40

<400> 3

cgggaagctg gggtcacttc ctgctccgtc ccagtcaccc tgaactggtc cccgggtcct 60 120 cagcetggaa teacceettg agaaggacee cagagteete ggegeecagt cegteeceeg aggcagggca ctgaaggggt taagtcccct ggggctggat tccgccttcc ggcttttccc 180 agaccccaga gccggtccct ggaacactgc agtcctgagc tctggg atg gag ccc 235 Met Glu Pro 1 gag act gcg ctg tgg ggc ccg gat ctg cag ggt ccg gaa cag agc ccc 283 Glu Thr Ala Leu Trp Gly Pro Asp Leu Gln Gly Pro Glu Gln Ser Pro 5 10 331 aac gat gct cac aga ggt gcc gag agt gaa aac gaa gag gag agc cct Asn Asp Ala His Arg Gly Ala Glu Ser Glu Asn Glu Glu Glu Ser Pro 20 25 379 cgg cag gaa agt tct ggg gag gag atc atc atg gga gac ccg gct cag

Arg Gln Glu Ser Ser Gly Glu Glu Ile Ile Met Gly Asp Pro Ala Gln

45

| | | | | | | ctg Leu | | | 427 |
|-----|-----|---|--|------|--|-------------------|--|--|------|
| | | | | | | cca Pro | | | 475 |
| | | | | | | cca Pro 95 | | | 523 |
| | | | | | | gag Glu | | | 571 |
| | | | | | | gtc Val | | | 619 |
| | | | | | | ccg Pro | | | 667 |
| | | | | | | ttc Phe | | | 715 |
| | | | | | | gag Glu 175 | | | 763 |
| | | | | | | tcg Ser | | | 811 |
| _ | - | _ | | | | aag Lys | | | 859 |
| | | | | | | cag Gln | | | 907 |
| | | | | | | tgc Cys | | | 955 |
| | Ser | | | | | tcc Ser 255 | | | 1003 |
| Pro | | | | | | ttc Phe | | | 1051 |

| | ctc Leu | | | | | | | | | | | | | | | 1099 |
|------------|-------------------|------------|------------|------------|------------|-------------------|------------|------------|------------|------------|-------------------|------------|------------|------------|------------|------|
| | ccc Pro | | | | | | | | | | | | | | | 1147 |
| | cag Gln | | | | | | | | | | | | | | | 1195 |
| ggg Gly | aag Lys 325 | agc Ser | ttc Phe | atc Ile | cag Gln | agc Ser 330 | tcg Ser | gag Glu | ctg Leu | acc Thr | cag Gln 335 | cac His | cag Gln | cgc Arg | acg Thr | 1243 |
| | aca Thr | | | | | | | | | | | | | | | 1291 |
| | cac His | | | | | | | | | | | | | | | 1339 |
| | ccg Pro | | | | | | | | | | | | | | | 1387 |
| _ | ttg Leu | _ | | | - | | | | | | | | | | | 1435 |
| | cca Pro 405 | | | | | | | | | | | | | | | 1483 |
| | cag Gln | | | | | | | | | | | | | | | 1531 |
| | aag Lys | | | | | | | | | | | | | | | 1579 |
| | acc Thr | | | | | | | | | | | | | | | 1627 |
| | cgc Arg | | | | | | | | | | | | | | | 1675 |
| _ | ccc Pro 485 | | | | | | | | | | | | | | | 1723 |
| aac | ctt | att | acc | cac | gtc | cgc | acg | cac | atg | gac | gag | aac | ctg | ttc | gtg | 1771 |

| Asn 500 | Leu | Ile | Thr | His | Val 505 | Arg | Thr | His | Met | Asp 510 | Glu | Asn | Leu | Phe | Val 515 | | |
|------------|-----|-----|-----|-----|-------------------|-----|-----|-----|-----|------------|-----|-----|-----|-----|------------|---|------|
| | | | | | aag Lys | | | | | | | | | | | 1 | L819 |
| | | | | | gag Glu | | | | | | | | | | | - | 1867 |
| | | | | | ggg Gly | | | | | | | | | | | - | 1915 |
| | | | | | cac His | | | | | | | | | | | - | 1963 |
| | | | | | atg Met 585 | | | | | | | | | | | : | 2011 |
| | | | | | gac Asp | | | | | | | | | | | : | 2059 |
| | | | | | cca Pro | | | | | | | | | | | : | 2107 |
| | | | | | aga Arg | | | | | | | | | | | : | 2155 |
| | | | | | ggc Gly | | | | | | | | | | | | 2203 |
| | | | | | tcc Ser 665 | | | | | | | | | | | | 2251 |
| | | | | | cag Gln | | | | | | | | | | | | 2299 |
| | | | | | att Ile | | | | | | | | | | | | 2347 |
| | | | Gly | | ccc Pro | | | Cys | | | | | | | | | 2395 |
| | | | | | ctg Leu | | | | | | | | | | | | 2443 |

735 725 730 2491 aag agc tcc cag aag agt cca gag ctg ggg aag agc tct tcc gtc ctc Lys Ser Ser Gln Lys Ser Pro Glu Leu Gly Lys Ser Ser Ser Val Leu 745 ctg gag cat ctc agg agc ccc ctg ggg gcc aga ccc tac cgc tgc tca 2539 Leu Glu His Leu Arg Ser Pro Leu Gly Ala Arg Pro Tyr Arg Cys Ser 760 765 2587 gat tgc agg gcc tcc ttc ctc gac cgc gtg gcc ctc acc cgg cac caa Asp Cys Arg Ala Ser Phe Leu Asp Arg Val Ala Leu Thr Arg His Gln 775 780 qaa acc cac acc cag gaa aaa ccc ccc aat ccc gag gac ccc cct cca 2635 Glu Thr His Thr Gln Glu Lys Pro Pro Asn Pro Glu Asp Pro Pro Pro 795 790 gag gca gtc acc ctg tcc aca gat cag gaa ggt gag ggc gag acc cct 2683 Glu Ala Val Thr Leu Ser Thr Asp Gln Glu Gly Glu Gly Glu Thr Pro 815 805 810 2731 acc ccc aca gag agc agc cat ggg gaa ggg caa aac ccc aaa acc Thr Pro Thr Glu Ser Ser His Gly Glu Gly Gln Asn Pro Lys Thr 825 830 2779 cta gtg gaa gaa aag ccc tat ctg tgc ccc gag tgt gga gcc ggc ttc Leu Val Glu Glu Lys Pro Tyr Leu Cys Pro Glu Cys Gly Ala Gly Phe 2827 aca gaa gtc gca gcc ctc ctg ctc cat agg agc tgc cac cca ggt gtc Thr Glu Val Ala Ala Leu Leu Leu His Arg Ser Cys His Pro Gly Val 2883 tcc ctq tqa aatqqqt ctqqaqacca qqqqcctcqc tctctccaqa qaqqaacact Ser Leu 870 2943 ggattttttc ccccaaaaaa attacatggg gaagggagga taaccctatc agatggtagt 3003 ggagtggagg agaaagaacc ctgggaaaaa atagtgcttt tacatcagtg atgagaaacc ctataaaatt gttggtggga agcacttata aggcagtaga gaaaaactgt tgggttagaa 3063 3123 accctataaa tatgtaggaa aaaaaaaagc cctttaagtc tgtagcagaa aaaccctata aaccatagtg gataaaagcc ctattaattg taggaagagg tcccaatatg tctctgaaca 3183 3243 accetataaa attgtatcaa aateettagg aaaateetat agggtegggg aagtacegca 3303

3363

3423

3483

atgqcactga cccgggagta acgaccctca gatgatgcag aaaacacctc tccctcactg

ttccacaatg ttctctccca gcagccctgg acagttctca tggaggaaga tgctcagctg

ggggagtggt gtgggaactg ggtggaactg aaaaggaaga aaataggaga aaaataagtg

accaggggct ctggagcccc cagaggagat aagacagaga ctaggaagaa aaggtcccac

3543 tcactgcact ggggtccttg gggtttgagg acaatgttag atgcattgtg tgttttaggg aaagaggcct atggaatccc ttatcaaaac cccaaaatgg ggaagaaaaa acaaattggc 3603 3663 aagccttttt aaaggctgaa gtttggggtg cagaaacatg gtcagtataa gaatactgag 3723 aaaccaccac aaaaattagg gggaagcagg gggcagtttc tgttagattg ctttgagggg 3783 agcaaaccaa ggtggggaaa ctgctcacat ggaacccagc taactccagg ccccttcttg 3843 cttccggaag cttgcgaaac tgcaagtagc aggcaggtcc cagctctcag agtgcaaggg 3903 ttctccccac accaggcagg accacccttt ctccctgccc agccagggac ttctactctg 3963 caaggcetee agggatggeg gttgceteee accaeattet caggteteae etteaggagg 4023 tttttcctqa aatctaactq ccqccagtgg cagctcattt cctcaaattc tgtcctcagt ggagagggag aaccgctgct agtcgacctc caaagaataa agcccttcag agacttaggg 4083 4143 cctgcgtgtg ggcctcccca gccttctcca agtcatcagg gccacagaat ttcttagaaa gctttgatta gccagcaact tgattttaaa cttcagctat tcatctttgt tggtagcaag 4203 gctgcccgc aggataaggc agccctcctc cctggtgggg atccagacac cccacagact 4263 4323 gagccacagg tecetgaaga gteegcagag geegtateac cetetteace etggtacetg 4383 gagccctcca gggttgtgag atccagccca atctcagcct tgagatggat gggatggatt 4443 ccactcactt ggctcccctc tcctccttgg ccagaaacct tcagaaggaa acagaaggat gagaaagcct cccttgccag gcccatctca cctcccacac cacacagtta cacattcata 4503 4563 aacagccaag cctgctcaca ggcctgggca tatgtggctg gggagagact gttgccctat 4623 gagatggtgg aagacagaaa actettgaat gaaacatetg taacaagact cecacteect 4683 aateetggag acceetteaa aaggagagga etetgeteag egetgggtea gaaaggaagt 4743 gcttgccaac cccagtaatg cgctgtcctg gagtaagagg actgagatgc cagctgtggg 4803 gcccgtcagg agtcacgtct caagagaggg gtgacccaga tttttgcaaa gctgtggaag 4863 agacctctaa ttccatgacc taaagtttca cacaaccttt atgcaaatcc ttcgccattt gtctatcctg gccccagcag gatgaccctc acatttaaaa ggatcaaaat ctcaaacatt 4923 4983 gtgggaggaa aacgccaggc ctgctgggga gagggagaag gcaggcctca ggattgtgcg 5043 gagaaccatg ttgccatctg cataacgccc cagcctatcc aagtgccagg aagggaagca 5103 ctactgaggg gggtgctgcc tgtgggtctc ttcctaacag aacgccaacg aagggggtgc 5163 cctgtaagaa tgagaagtct gccaagtcta aggagatgtc cccctcccca agatggtagc

5223 agaaactgaa gttgtaaaga caaacctctc gcaaagccag aagcagatgc tctcaaaaac 5283 aggectaagg aatctgttcc cagaagaggc ctggggagca gettttgggg ctcagaggec 5343 agtctggggg gaggcaggca gtgtgactga aggccggcgc aggcctgttt cggactaggg 5403 tqtaqcatat ttagtaggat gctcacatgt tgtagggatg ggtgagggct aaggggcttg 5463 gctcttgtac gggatggatg atgtgagcgc acagttcgta gggtcaggcg tcttgcggta ggaaaggggc agggttcgat acctctgggg tgtggcgtgt tctctcttgc ccacctcccc 5523 5583 aagctccttt aaccaggagc ttccctccca gagtgagcta cgtcttggga ggattgaagc 5643 agggaggtgg cagtggctag agttgctcag atcggctcag atgaaggccc aggggtcaaa 5703 gcatttgctc accacaactg agtctggaag gcactgtgca gccaaaccct cttgtccaat 5763 cacatgcagg ccccaggccc ctcagaaacg ccctggtgga gggaggagcc caacagatag 5823 qqtcaqqqaq tcaqtqqqqa aagcagaggg gagagcttag agtgaaactt gaaactgcta 5883 tggacatgga ggtcagatgg gaacttggaa ctgggcatga atcctgaatg gtggggaaac 5943 ttgaagccat cacagccaag aggtggggtt ccatgacctg ccctgaggtc agcagctccc 6003 agactccttt gccctcacta gggaagcccc agtcaccgga cttgtcccca tggagtgaag 6063 agaggcccca tttttgagtg ttgtgtgtcc aaaacagtgt tgtgtcactg gtggtcatgt 6109 tgcttagttg aagagaataa agggaataag atttaaaaaa aaaaaa

<210> 4

<211> 3657

<212> DNA

<213> Homo sapiens

<220>

<221> CDS

<222> (645)..(3254)

<400> 4

acggctaggt acgaggcctg gtgttagcag agttgagtcc actccgcacc ccaccetctg 60
tctggtacag cttaccaaac caaagtgccc aaagccgtga catcccggcc ggcggctcgc 120
aggcccccgc cctccgcacg tcacggccgc cgggtgcagt gcccctagg ggcccctggg 180
acgaggagga agcgccaggt ccttcccgcc gccgccgccg ccgccgccgc cgcctgctcc 240
cctggcacgc gcccgccgc cctcggcagc cgcagctccg tgtcccctga gaaccagccg 300
tcccgcgcca tgggcacgcg tctgccgctc gtcctgcgcc agctccgccg cccgccccag 360

420 ccccqqqcc ctccqcqccg cctccgtgtg ccctgtcgcg ctagcagcag gcgcgcggcg gcaggcggct ggcgtgtggc cgggagggcc tgcttggaca gcggcggccg caggatggcc 480 540 aggcccqqaq cagctgcagc cccggcggcc gaacgcccgc ggcgcgggac tccatcgtca gagaagtcat tcagaattca aaagaagttc taagtttatt gcaagaaaaa aaccctgcct 600 tcaagccggt tcttgcaatt atccaggcag gtgacgacaa cttg atg cag gaa atc 656 Met Gln Glu Ile 704 aac cag aat ttg gct gag gag gct ggt ctg aac atc act cac att tgc Asn Gln Asn Leu Ala Glu Glu Ala Gly Leu Asn Ile Thr His Ile Cys 10 15 5 ctc cct cca gat agc agt gaa gcc gag att ata gat gaa atc tta aag 752 Leu Pro Pro Asp Ser Ser Glu Ala Glu Ile Ile Asp Glu Ile Leu Lys 25 800 atc aat gaa gat acc aga gta cat ggc ctt gcc ctt cag atc tct gag Ile Asn Glu Asp Thr Arg Val His Gly Leu Ala Leu Gln Ile Ser Glu 40 848 aac ttg ttt agc aac aaa gtc ctc aat gcc ttg aaa cca gaa aaa gat Asn Leu Phe Ser Asn Lys Val Leu Asn Ala Leu Lys Pro Glu Lys Asp 60 896 gtg gat gga gta aca gac ata aac ctg ggg aag ctg gtg cga ggg gat Val Asp Gly Val Thr Asp Ile Asn Leu Gly Lys Leu Val Arg Gly Asp 944 gcc cat gaa tgt ttt gtt tca cct gtt gcc aaa gct gta att gaa ctt Ala His Glu Cys Phe Val Ser Pro Val Ala Lys Ala Val Ile Glu Leu 90 ctt gaa aaa tca gta ggt gtc aac cta gat gga aag aat ttg gta 992 Leu Glu Lys Ser Val Gly Val Asn Leu Asp Gly Lys Lys Ile Leu Val 110 105 1040 gtg ggg gcc cat ggg tct ttg gaa gct gct cta caa tgc ctg ttc cag Val Gly Ala His Gly Ser Leu Glu Ala Ala Leu Gln Cys Leu Phe Gln 120 130 aga aaa ggg tcc atg aca atg agc atc cag tgg aaa aca cgc cag ctt 1088 Arg Lys Gly Ser Met Thr Met Ser Ile Gln Trp Lys Thr Arg Gln Leu 135 140 1136 caa agc aag ctt cac gag gct gac att gtg gtc cta ggc tca cct aag Gln Ser Lys Leu His Glu Ala Asp Ile Val Val Leu Gly Ser Pro Lys 150 155 1184 cca qaa qaq att ccc ctt act tgg ata caa cca gga act act gtt ctc Pro Glu Glu Ile Pro Leu Thr Trp Ile Gln Pro Gly Thr Thr Val Leu 175 170 1232

aac tgc tcc cat gac ttc ctg tca ggg aag gtt ggg tgt ggc tct cca

| Asn | Cys | Ser | His | Asp 185 | Phe | Leu | Ser | Gly | Lys 190 | Val | Gly | Cys | Gly | Ser 195 | Pro | | |
|-----|-----|-----|-----|------------|-----|-----|-----|-----|------------|-----|-----|-----|-----|-------------------|-----|---|------|
| | | | | | | | | | | | | | | ctc Leu | | • | 1280 |
| | | | | | | | | | | | | | | aga Arg | | | 1328 |
| | _ | - | _ | _ | | | | | _ | | | | | aaa Lys | | | 1376 |
| - | | | | | | | _ | _ | | | | | _ | gga Gly | | | 1424 |
| | | | _ | | _ | ~ | | _ | _ | | | | _ | ctt Leu 275 | | | 1472 |
| | 9 | | _ | | | - | | _ | | ~ | | ~ | ~ | ttg Leu | | | 1520 |
| | | - | | | _ | - | | _ | - | | | | - | tta Leu | - | | 1568 |
| - | | | | | | | | | - | | _ | _ | | gtc Val | | | 1616 |
| | | | ~ ~ | _ | - | _ | | - | | _ | | - | | tcc Ser | | | 1664 |
| | | | | | | | | | | | | | | aaa Lys 355 | | | 1712 |
| | - | | | | | | _ | _ | _ | | | _ | | gag Glu | | | 1760 |
| | | | _ | | | - | | | _ | | | - | _ | aat Asn | | | 1808 |
| _ | - | | _ | _ | | _ | _ | | | | | | | acg Thr | | | 1856 |
| | | | | | | | | | | | | | | ggt Gly | | | 1904 |

| 405 | | | 410 | | | | | 415 | | | | 420 | |
|-----|-------|-------------------|-----|---|---|---|---|-----|---|---|---|-----|------|
| | | gaa Glu 425 | | | | | | | | | | | 1952 |
| | | ccg Pro | | | | | | | | | | | 2000 |
| | | atc Ile | | | | | | | | | | | 2048 |
| | | cga Arg | | | | | | | | | | | 2096 |
| | _ | cat His | | | _ | | _ | | - | _ | - | _ | 2144 |
| | | gcg Ala 505 | | | | | | | | | | | 2192 |
| | | gga Gly | | | | | | | | | | | 2240 |
| | | gat Asp | | | | | | | | | | | 2288 |
| | | ata Ile | | | | | | | | | | | 2336 |
| | | cat His | | | | | | | | | | | 2384 |
| | | gct Ala 585 | | | | | | | | | | | 2432 |
| | | acc Thr | | | | | | | | | | | 2480 |
| | | atc Ile | _ | _ | - | _ | | | | | | | 2528 |
| | | gca Ala | | | | | | | | | | | 2576 |

| | | | | | | | | | gaa Glu | | | 2624 |
|---|---|-------|---|---|---|---|-----|-------|-------------------|---|--|------|
| | | | | | | | | | ctc Leu | | | 2672 |
| | | | | | | | | | gtg Val | | | 2720 |
| | - | _ | | _ | - | | | | gtg Val | | | 2768 |
| | | | | | | | | | tat Tyr 720 | | | 2816 |
| | | | | | | | | | gtg Val | | | 2864 |
| | | | | | | | | | gtt Val | | | 2912 |
| | | | | | | | | | tat Tyr | | | 2960 |
| | | | | | | | | | gat Asp | | | 3008 |
| | | | | | | | | | aag Lys 800 | | | 3056 |
| | | | | _ | | | ~ ~ | | agg Arg | - | | 3104 |
| | | | | | | | | | Gly | | | 3152 |
| - | - | _ | - | _ | | _ | | - | ccc Pro | | | 3200 |
| | | | | | | | | | caa Gln | | | 3248 |

| ttc taa gtg Phe * 870 | gacaagg cto | ctcacagg acc | ccgatgca gad | stootgaa aca | agactact | 3304 |
|-----------------------------|-------------|--------------|--------------|--------------|------------|------|
| ctttgccttt | ttgctgcagt | tggagaagaa | actgaatttg | aaaaatgtct | gttatgcaat | 3364 |
| gctggagaca | tggtgaaata | ggccaaagat | ttcttcttcg | ttcaagatga | attctgttca | 3424 |
| cagtggagta | tggtgttcgg | caaaaggacc | tccaccaaga | ctgaaagaaa | ctaatttatt | 3484 |
| tctgtttctg | tggagtttcc | attatttcta | ctgcttacac | tttagaatgt | ttattttatg | 3544 |
| gggactaagg | gattaggagt | gtgaactaaa | aggtaacatt | ttccactctc | aagttttcta | 3604 |
| ctttgtcttt | gaactgaaaa | taaacatgga | tctagaaaac | caaaaaaaaa | aaa | 3657 |
| | | | | | | |
| <210> | 5 | | | | | |

<210> 5
<211> 806
<212> DNA
<213> Homo sapiens
<220>
<221> CDS
<222> (341)..(487)

<400> 5 60 accacgaggt cacggtaagg gtggaaccag gattctgccc ggatggggat ctggcatggg 120 ggggtgggg gttgatttgg gaagcagagc acagcagccc aaatttgctt gtaatgtcgg 180 cggctacaga ggaaaggcca agaatggaca gattaataaa gggccatcta caagccacga ataaggccat caccagaagc caaccccgcc agtccttgat ctaggacttc cagacttcaa 240 300 aacttttcta geetgetgge etgeetteae tgteetgggg gagaettgga gagaeeaggt 355 ggactggagt agactgttga gagacgctgg tctggtgaag atg tcc agg aaa cca Met Ser Arg Lys Pro 1 403 cga gcc tcc agc cca ttg tcc aac aac cac cca cca aca cca aag agg

cga gga agt gga agg ttc cca aga caa ccc gga agg gaa aag gga ccc
Arg Gly Ser Gly Arg Phe Pro Arg Gln Pro Gly Arg Glu Lys Gly Pro
25 30 35

atc aag gaa gtt cca gga aca aaa ggc tct ccc taa aaga ccgccgcttc
Ile Lys Glu Val Pro Gly Thr Lys Gly Ser Pro *

Arg Ala Ser Ser Pro Leu Ser Asn Asn His Pro Pro Thr Pro Lys Arg

aaaaaaacct gaggaatgga gtgggccaac actatccagc cactctgacc agccgaacga 561

ggaactcaat caaaatgagc catagcggga ccacaagggc aaggagacca ccaccttctc 621
cagtctctct tcggacagcc agtaattccc gggcaaggcc agagacttca agtctatctg 681
aaaagtctcc agaggtctaa ccccagataa atagccaaca gggtgtagag tacattttac 741
accccaaaga gtgtgcccca tggtgatgaa aataaagtga acatgttgca aactgaaaaa 801
aaaaa

<210> 6 <211> 842 <212> DNA <213> Homo sapiens <220> <221> CDS <222> (341)..(523)

<400> 6

accacgaggt cacggtaagg gtggaaccag gattctgcc ggatggggat ctggcatggg 60 ggggtggggg gttgatttgg gaagcagagc acagcagcc aaatttgctt gtaatgtcgg 120 cggctacaga ggaaaggcca agaatggaca gattaataaa gggccatcta caagccacga 180 ataaggccat caccagaagc caaccccgcc agtccttgat ctaggacttc cagacttcaa 240 aacttttcta gcctgctggc ctgccttcac tgtcctgggg gagacttgga gagaccaggt 300 ggactggagt agactgttga gagacgctgg tctggtgaag atg tcc agg aaa cca 355

cga gcc tcc agc cca ttg tcc aac aac cac cca cca aca cca aag agg 403
Arg Ala Ser Ser Pro Leu Ser Asn Asn His Pro Pro Thr Pro Lys Arg
10 15 20

Met Ser Arg Lys Pro

5

1

cga gga agt gga agg cat cct ctc atc cct ggc cca gaa gcc cta tca

Arg Gly Ser Gly Arg His Pro Leu Ile Pro Gly Pro Glu Ala Leu Ser

25 30 35

aag ttc cca aga caa ccc gga agg gaa aag gga ccc atc aag gaa gtt
Lys Phe Pro Arg Gln Pro Gly Arg Glu Lys Gly Pro Ile Lys Glu Val
40
45
50

cca gga aca aaa ggc tct ccc taa aagaccgccg cttcaaaaaa acctgaggaa 553
Pro Gly Thr Lys Gly Ser Pro *
55 60

tggagtgggc caacactatc cagccactct gaccagccga acgaggaact caatcaaaat 613 gagccatagc gggaccacaa gggcaaggag accaccacct tctccagtct ctcttcggac 673

| agccagtaat tecegggeaa ggecagagae tteaagteta tetgaaaagt etecagaggt | 733 |
|--|-----|
| ctaaccccag ataaatagcc aacagggtgt agagtacatt ttacacccca aagagtgtgc | 793 |
| cccatggtga tgaaaataaa gtgaacatgt tgcaaactga aaaaaaaaa | 842 |
| <210> 7 <211> 1805 <212> DNA <213> Homo sapiens <220> <221> CDS <222> (9)(1517) | |
| <400> 7 | |
| actgggaa atg tgc agg att ctg aat gat gag gct gcc tgg aat gag ttg Met Cys Arg Ile Leu Asn Asp Glu Ala Ala Trp Asn Glu Leu 1 5 10 | 50 |
| gcc aga gct tgt cta cat cac atg gaa gtg gag ttt gca atc cgt gtt Ala Arg Ala Cys Leu His His Met Glu Val Glu Phe Ala Ile Arg Val 15 20 25 30 | 98 |
| tat cgg aga att gga aat gtt ggc ata gtg atg tcc ttg gaa caa ata Tyr Arg Arg Ile Gly Asn Val Gly Ile Val Met Ser Leu Glu Gln Ile 35 40 45 | 146 |
| aag gga ata gag gac tac aat ctt ttg gca gga cac ctt gcc atg ttt Lys Gly Ile Glu Asp Tyr Asn Leu Leu Ala Gly His Leu Ala Met Phe 50 55 60 | 194 |
| acc aac gat tat aac ctg gct cag gac ttg tac ctt gca tcc agc tgt Thr Asn Asp Tyr Asn Leu Ala Gln Asp Leu Tyr Leu Ala Ser Ser Cys 65 70 75 | 242 |
| cct att gct gcc ctg gag atg aga agg gat tta cag cat tgg gac agt Pro Ile Ala Ala Leu Glu Met Arg Arg Asp Leu Gln His Trp Asp Ser 80 85 90 | 290 |
| gct cta caa ctg gca aag cat ttg gcc cca gac cag ata cct ttt ata Ala Leu Gln Leu Ala Lys His Leu Ala Pro Asp Gln Ile Pro Phe Ile 95 100 105 110 | 338 |
| tca aaa gaa tat gct att cag ctt gaa ttc gcg ggt gat tat gta aat Ser Lys Glu Tyr Ala Ile Gln Leu Glu Phe Ala Gly Asp Tyr Val Asn 115 120 125 | 386 |
| gct ttg gct cat tat gag aaa gga ata aca ggt gat aat aag gaa cat Ala Leu Ala His Tyr Glu Lys Gly Ile Thr Gly Asp Asn Lys Glu His 130 135 140 | 434 |
| gat gaa gct tgt ctg gct gga gtg gcc cag atg tcc ata aga atg gga Asp Glu Ala Cys Leu Ala Gly Val Ala Gln Met Ser Ile Arg Met Gly | 482 |

| | 145 | | 1 | L50 | | | 155 | | | | |
|---------------------------|-----|---------------------------|-------|-----|-----|---|-------|---|---|---|------|
| gac ata Asp Ile 160 | | | | | | | | | | | 530 |
| ctt aaa Leu Lys 175 | | | | | | | | | | | 578 |
| gaa gcg Glu Ala | | | | | | | | | | | 626 |
| tct gtt Ser Val | | | | | | | | | | | 674 |
| ccc cac Pro His | _ | | Lys I | | _ | - | | | | | 722 |
| | | aga tac Arg Tyr | | | | | | | | | 770 |
| aaa cag Lys Gln 255 | | agt gta Ser Val 260 | | | | | | | | | 818 |
| | | gtc aat Val Asn 275 | | | | | | | | | 866 |
| - | | gcc agg Ala Arg | | | _ | | - | | | | 914 |
| Ala Ile | - | ctt gtc Leu Val | Met S | | _ | | _ | _ | | | 962 |
| | - | cac aac His Asn | | | | | | | | | 1010 |
| _ | _ | act aat Thr Asn 340 | - | - | | - | _ | | | | 1058 |
| | | aga tat Arg Tyr 355 | | | | | | _ | _ | _ | 1106 |
| | | cga gca Arg Ala | | | Phe | _ | _ | | - | _ | 1154 |

| gaa gat aat gtg gca ata gaa atg gca att gaa act gtt ggt cag gcc Glu Asp Asn Val Ala Ile Glu Met Ala Ile Glu Thr Val Gly Gln Ala 385 390 395 | 1202 |
|---|------|
| aaa gat gaa ctg ctg acc aat cag ctg ata gac cat ctc ctg ggg gag Lys Asp Glu Leu Leu Thr Asn Gln Leu Ile Asp His Leu Leu Gly Glu 400 405 410 | 1250 |
| aac gat agc atg cct aag gat gcc aag tac ctg ttc cgc ttg tac atgAsn Asp Ser Met Pro Lys Asp Ala Lys Tyr Leu Phe Arg Leu Tyr Met415420 | 1298 |
| gct ctg aag caa tac cga gaa gct gcc cag act gcc atc atc att gcc Ala Leu Lys Gln Tyr Arg Glu Ala Ala Gln Thr Ala Ile Ile Ile Ala 435 440 445 | 1346 |
| aga gaa gag cag tct gca ggc aac tac cgg aat gca cac gat gtt ctc Arg Glu Glu Gln Ser Ala Gly Asn Tyr Arg Asn Ala His Asp Val Leu 450 455 460 | 1394 |
| ttc agt atg tat gca gaa ctg aaa tcc cag aag atc aaa att ccc tcc Phe Ser Met Tyr Ala Glu Leu Lys Ser Gln Lys Ile Lys Ile Pro Ser 465 470 475 | 1442 |
| gag atg gcc acc aac ctc atg att ctg cac agc tat ata cta gta aga Glu Met Ala Thr Asn Leu Met Ile Leu His Ser Tyr Ile Leu Val Arg 480 485 490 | 1490 |
| ttc atg tta aaa atg gag atc aca tga aaggg gctcgcatgc tcattcgggt Phe Met Leu Lys Met Glu Ile Thr * 495 500 | 1542 |
| ggccaacaac atcagcaaat ttccatcaca cattgtaccc atcctgacgt caactgtgat | 1602 |
| tgagtgtcac agggcaggcc tgaagaactc tgctttcagc ttcgcagcta tgttgatgag | 1662 |
| gcctgaatac cgcagcaaaa tagatgccaa atacaaaaag aagatcgagg gaatggtcag | 1722 |
| gagacccgat atatctgaga tagaagaggc cacgactcca tgtccattct gcaaatttct | 1782 |
| tctcccagag agagaactcc tca | 1805 |

```
<210> 8 <211> 1523
```

<212> DNA

<213> Homo sapiens

<220>

<221> CDS

<222> (231)..(1202)

<400> 8

tattttgaag ctgtttaccc tcgcagctct ctgactggca cccctgcctg cctgcccggc 60

| cctgcacaac atgca | ageeet eeggee | ctcga gggtcccgg | gc acgtttggtc gg | tggcctct 120 |
|--|-----------------------|-----------------|---|-----------------------------|
| gctgagtctg ctgct | cetge tgetge | ctgct ccagcctgt | ta acctgtgcct ac | eaccacgcc 180 |
| aggcccccc cagag | geeete accac <u>q</u> | getgg gegeeeea | | atg ccg 236 Met Pro 1 |
| ggc acc tac gct Gly Thr Tyr Ala 5 | | | | |
| ctg caa gag cag Leu Gln Glu Gln 20 | | | | |
| | | Val Leu Arg G | ag gtt tac cag a ln Val Tyr Gln I 45 | |
| | | | ac ggc cag acc a yr Gly Gln Thr S | |
| | | | ag ttc tgg tca g ln Phe Trp Ser A 80 | |
| | | | tg cgc ctc acc c eu Arg Leu Thr I 95 | |
| | | | cc tat tct gag c er Tyr Ser Glu I 110 | |
| | | Leu Asn Asp T | ct cag aaa ttg g hr Gln Lys Leu <i>A</i> 25 | |
| | | | ac aat agc ctc t sp Asn Ser Leu S 1 | |
| | | | ac ctg acg ctc a yr Leu Thr Leu 1 160 | |
| _ | | | ct aag ggc gtc o la Lys Gly Val F 175 | |
| | | Leu Thr Asp P | tt ggt gag aag g he Gly Glu Lys V 190 | |

| gca gaa atg aac cgc ctg ggc atg atg gta gac tta tcc cat gtc tca Ala Glu Met Asn Arg Leu Gly Met Met Val Asp Leu Ser His Val Ser 195 200 205 210 | 860 |
|---|------|
| gat gct gtg gca cgg cgg gcc ctg gaa gtg tca cag gca cct gtg atc Asp Ala Val Ala Arg Arg Ala Leu Glu Val Ser Gln Ala Pro Val Ile 215 220 225 | 908 |
| ttc tcc cac tcg gct gcc cgg ggt gtg tgc aac agt gct cgg aat gtt Phe Ser His Ser Ala Ala Arg Gly Val Cys Asn Ser Ala Arg Asn Val 230 235 240 | 956 |
| cct gat gac atc ctg cag ctt ctg aag aag aac ggt ggc gtc gtg atg Pro Asp Asp Ile Leu Gln Leu Leu Lys Lys Asn Gly Gly Val Val Met 245 250 255 | 1004 |
| gtg tct ttg tcc atg gga gta ata cag tgc aac cca tca gcc aat gtg Val Ser Leu Ser Met Gly Val Ile Gln Cys Asn Pro Ser Ala Asn Val 260 265 270 | 1052 |
| tcc act gtg gca gat cac ttc gac cac atc aag gct gtc att gga tcc Ser Thr Val Ala Asp His Phe Asp His Ile Lys Ala Val Ile Gly Ser 275 280 285 290 | 1100 |
| aag ttc atc ggg att ggt gga gat tat gat ggg gcc ggc aag tac agg Lys Phe Ile Gly Ile Gly Gly Asp Tyr Asp Gly Ala Gly Lys Tyr Arg 295 300 305 | 1148 |
| aag aaa aca aag tgc aaa gcc cct tgg agg aca agt tcc cgg atg agc Lys Lys Thr Lys Cys Lys Ala Pro Trp Arg Thr Ser Ser Arg Met Ser 310 315 320 | 1196 |
| agc tga gcagtteetg ecaeteegae eteteaegte tgegteagag acagagtetg Ser * | 1252 |
| acttcaggcc aggaactcac tgagattccc atacactgga cagccaagtt accagccaag | 1312 |
| tggtcagtct cagagtcctc cccccacatg gccccagtcc ttgcagttgt ggccaccttc | 1372 |
| ccagtcctta ttctgtggct ctgatgaccc agttagtcct gccagatgtc actgtagcaa | 1432 |
| gccacagaca ccccacaaag ttcccctgtt gtgcaggcac aaatatttcc tgaaataaat | 1492 |
| gttttggaca tagaaacaga aaaaaaaaaa a | 1523 |

```
<210> 9
```

<211> 1608

<212> DNA

<213> Homo sapiens

<220>

<221> CDS

<222> (49)..(1608)

| cga | | 1ac c | | ıccaa | ac ac | tgga | ıgaag | g aag | gaaaa | aac | ctca | itttg | | : Ala | gaa Glu | 57 |
|-----|-------------------|-------|---|-------|-------|------|-------|-------|-------|-----|------|-------|---|-------|------------|-----|
| | gaa Glu 5 | | | | | | | | | | | | | | | 105 |
| | acc Thr | | | | | | | | | | | | | | | 153 |
| | aag Lys | | _ | | | | | | | | | | | | | 201 |
| | agg Arg | | | | | | | | | | | | | | | 249 |
| ~ ~ | cag Gln | _ | | | | | | | | | | | | | | 297 |
| | tgt Cys 85 | | | | | | | | | | | | | | | 345 |
| | ctg Leu | | | | | | | | | | | | | | | 393 |
| | aag Lys | | | - | | | | | | | | | | | | 441 |
| _ | cat His | ~ | | | | - | Lys | | - | | | | _ | | | 489 |
| | ata Ile | | | | | | | | | | | | | | | 537 |
| | ata Ile 165 | | - | | | | | | | | | | | | | 585 |
| | tgt Cys | | | | | | | | | | | | | | | 633 |
| | att Ile | | | | | | | | | | | | | | | 681 |

| | - | | - | aat Asn | _ | | | | | _ | | | | | 729 |
|---|---|---|---|-------------------|---|---|---|---|---|---|---|---|---|---|----------|
| | | | | tcc Ser | | | | | | | | | | | 777 |
| | | | | aaa Lys | | | | | | | | | | | 825 |
| | | | | gac Asp | | | | | | | | | | | 873 |
| _ | | | | att Ile 280 | | _ | | _ | | | - | | | _ | 921 |
| | | | | ctc Leu | | | | | | | | | | | 969 |
| | | | | tta Leu | | | | | | | | | | | 1017 |
| - | | | | gtg Val | | | | | | | | | | | 1065 |
| | | | | tta Leu | | | | | | | | | | | 1113 |
| | - | | | gga Gly 360 | | | | | - | - | | | | | 1161 |
| | | | _ | gct Ala | | _ | - | | | | - | | | - | 1209 |
| | _ | _ | ~ | tac Tyr | | - | _ | - | | | | | | | 1257 |
| _ | | _ | | gaa Glu | - | - | | _ | | _ | _ | _ | | | 1305 |
| _ | | _ | _ | ttt Phe | | | | | | | | | _ | | 1353 |

| ccc aag tta cgg agc agg cac acg cct cac aag aca ctc atg ccc tac Pro Lys Leu Arg Ser Arg His Thr Pro His Lys Thr Leu Met Pro Tyr 440 445 450 | 1401 | | | | | | | | |
|---|------|--|--|--|--|--|--|--|--|
| gcg tcc ctc ttc cag tcg cac tcc tgc aag acc aag acc tcc ccg tgt Ala Ser Leu Phe Gln Ser His Ser Cys Lys Thr Lys Thr Ser Pro Cys 455 460 465 | 1449 | | | | | | | | |
| gtc ctg tca gac cgt ggc aaa gct cca gat ccc caa gca ggc aac ttt Val Leu Ser Asp Arg Gly Lys Ala Pro Asp Pro Gln Ala Gly Asn Phe 470 475 480 | 1497 | | | | | | | | |
| gtt ctt gtt ttt cct ttc aat gaa gca act ctc gga gct tcc agg aat Val Leu Val Phe Pro Phe Asn Glu Ala Thr Leu Gly Ala Ser Arg Asn 485 490 495 | 1545 | | | | | | | | |
| gga tta aat gtc aaa aga ata atc caa gag ctc cag aaa cta atg aat Gly Leu Asn Val Lys Arg Ile Ile Gln Glu Leu Gln Lys Leu Met Asn 500 505 510 515 | 1593 | | | | | | | | |
| aag caa cat tcc taa Lys Gln His Ser * 520 | 1608 | | | | | | | | |
| <210> 10 <211> 1891 | | | | | | | | | |
| <212> DNA <213> Homo sapiens | | | | | | | | | |
| <220> <221> CDS <222> (362)(1084) | | | | | | | | | |
| <220> <221> misc_feature <222> (1)(1891) <223> n = a,t,c or g | | | | | | | | | |
| <400> 10 ctggtggaat tcgaagcaaa gggtgtagct gatgtgcaga ttgggattga atgtgggaat | 60 | | | | | | | | |
| gagggtgtga tgggagtgat tggaaatatt gcagccctgc atagtctcca tcagggatgt | 120 | | | | | | | | |
| gacaaagtgg ataateteta eeacgtgaga aactteeaae attaettgea aateagattt | 180 | | | | | | | | |
| aatgaataaa ataaagctgt agcacttggc acattcattg ggacccttac ccaaacatta | 240 | | | | | | | | |
| tcaatattgt gtacgttatc tttattatca ggtcacaaaa gatgtcataa aagaatttgc | 300 | | | | | | | | |
| agatgacggc gtcaagtacc tggaactaag gagcacaccc agaagagaaa atgctactgg | 360 | | | | | | | | |
| a atg act aaa aag act tat gtg gaa tct ata ctt gaa ggt ata aaa Met Thr Lys Lys Thr Tyr Val Glu Ser Ile Leu Glu Gly Ile Lys | 406 | | | | | | | | |

| 1 | | 5 | | | 10 | | | | | 15 | |
|---------------------------------------|---------------------------|---------------------------|------------------|-----------------------|-------------------|-------------------|------------|------------|-------------------|-------------------|------|
| cag tcc aa Gln Ser Ly | | | | | | | | | | | 454 |
| gtt gac ag Val Asp Ar | | | Leu Va | | | | | | | | 502 |
| gcc gag gad Ala Glu Gl 5 | ı Phe Phe | | | | | | | | | | 550 |
| ctc agt gg Leu Ser Gl 65 | | | | | | | | | | | 598 |
| ctt tta ga Leu Leu Gl 80 | | | | | | | | | | | 646 |
| gag att cc Glu Ile Pro | | | | | | | | | | | 694 |
| cct gac ag Pro Asp Ar | | | | e Leu | | | | | | | 742 |
| tcc ctg ga Ser Leu Asp 13 | Leu Val | | | | | | | | | | 790 |
| ctc tgt tte Leu Cys Len 145 | g acc tca ı Thr Ser | aac gtc Asn Val 150 | aaa ag Lys Se | t cag r Gln | Thr | gtt Val 155 | cca Pro | tct Ser | tat Tyr | gac Asp | 838 |
| cag cac cat Gln His Hi: 160 | ttc gga S Phe Gly | ttc tgg Phe Trp 165 | tac ag Tyr Se | c att r Ile | gcc Ala 170 | cat His | cct Pro | tct Ser | gtg Val | atc Ile 175 | 886 |
| tgt act ga Cys Thr As _l | gat aag Asp Lys 180 | ggt gtt Gly Val | ttt gc Phe Al | a aca a Thr 185 | cac His | ctt Leu | tct Ser | caa Gln | gag Glu 190 | tac Tyr | 934 |
| cag ctg gca Gln Leu Ala | | | | u Thr | | | | | | | 982 |
| ctg tct ta Leu Ser Ty: 210 | Glu Ser | | | | | Ser | | | | | 1030 |
| tct gaa cto Ser Glu Leo 225 | | | | | Lys : | | | | | | 1078 |

| att taa gct Ile * 240 | tataatga ggt | tgaactac tto | ctgagtat gtg | gtttcaat caa | agttcctg | 1134 |
|-----------------------------|--------------|--------------|--------------|--------------|------------|---------|
| ccatatccca | cttagtaaaa | cagtccacca | ctcctttgaa | gcatagcaac | caagttcctt | 1194 |
| gggctctatc | accagcacct | tacacatggc | aggtactcag | taaatacgtg | tcttcaactg | 1254 |
| actcacaagc | tctcaggtgc | ttactgggtg | ggacttgact | gttgttgcta | attaaatccc | 1314 |
| cattccacca | gtgattattg | tgactcagca | gtccttccct | attagtgatc | ataaaatttc | 1374 |
| agggaaatcg | aagtttctca | tcaggaaatg | ttttggaatt | actagtataa | agttaggaaa | 1434 |
| gtggggaaat | taggttactg | ccgagacctt | taagccttct | aaacagcttt | atattttatt | 1494 |
| gtgcatactt | taatcagact | cccttcactc | gctttaagtt | tttaaaagta | ttccccagcc | 1554 |
| ggatgtgatg | gctcatgcct | gtaatcccag | cactttggga | agccaaagtg | ggcagattgc | 1614 |
| ttgatcctag | gagttcagta | gcagcctagg | caacatggag | aaaccctgtc | tctacaaaaa | 1674 |
| caaaaaaaca | aaaaaccgga | aattagtcag | gcacggtggt | acacacctgt | agtcccagcc | 1734 |
| accagggagg | ctaaggtggg | aggagacctg | atcccagggg | atgtttgagg | ctgcagtgag | 1794 |
| ctggagtgca | gtgacatgat | cacagatcac | tgcagctttc | agttttaaaa | cagcttttat | 1854 |
| tacattntct | ttataaaaa | ctdatttcta | ccttada | | | 1 2 0 1 |